

# ESTIMATING THE NEIGHBORHOOD INFLUENCE ON DECISION MAKERS: THEORY AND AN APPLICATION ON THE ANALYSIS OF INNOVATION DECISIONS

NIKOLAUS HAUTSCH AND STEFAN KLOTZ

NIKOLAUS.HAUTSCH@UNI-KONSTANZ.DE    STEFAN.KLOTZ@UNI-KONSTANZ.DE

CENTER OF FINANCE AND ECONOMETRICS (COFE)  
UNIVERSITY OF KONSTANZ

**ABSTRACT.** When making decisions, agents tend to make use of decisions others have made in similar situations. Ignoring this behavior in empirical models can be interpreted as a problem of omitted variables and may seriously bias parameter estimates and harm inference. We suggest a possibility of integrating such outside influences into models of discrete choice decisions by defining an abstract space in which agents with similar characteristics are neighbors who possibly influence each other. In order to correct for correlations between the characteristics, the design of this space allows for nonorthogonality of its dimensions. Several Monte Carlo simulations show the small sample properties of spatial models with binary choice. When applying the estimator to innovation decisions data of German firms, we find evidence for the existence of neighborhood effects.

*JEL classification:* C21, C25, C44, O33, R15

*Key Words:* decision models; discrete choice; neighborhood influence; spatial econometrics; social space; Euclidean measure

*An earlier version of this papers was presented at the 1999 meeting of the Econometric Society in Santiago de Compostela, Spain. We would like to thank the ZEW for providing us with the Mannheim Innovation Panel. Financial support by the CoFE and the Deutsche Forschungsgemeinschaft is gratefully acknowledged. For helpful comments, we would like to thank Frank Gerhard, Dieter Hess, Joachim Inkmann, Michael Lechner, Winfried Pohlmeier, participants of workshops at the Universities of Konstanz and St. Gallen, and an anonymous referee. The usual disclaimer applies.*

*Corresponding author: Nikolaus Hautsch, Box D 124, University of Konstanz, 78457 Konstanz, Germany.*

---

May 18, 2001.

## 1. INTRODUCTION

The amount of information is a key factor in decision making processes. Therefore, individuals try to take as much information into account as possible. So it may be sensible, when uncertainty exists, also to use information about the decisions of others, who have had a comparable decision to make, as a yardstick. Furthermore, from a sociological point of view, humans often tend to seek assurance as to whether their thinking is reasonable. As a result, an individual's decision for a particular alternative is more likely when he realizes that people with similar features have come to the same conclusion. An additional rationale in observing the decisions of others occurs when one is held responsible for the success of his decision - as e.g. managers are: a mistake can be more readily excused if others have also made the same mistake. For all of these reasons, we might expect that the decisions of agents in similar situations should correlate even after controlling for similarity. This paper aims to show how this interdependence of decisions can be modeled and estimated.

The benefit of integrating such outside influences is twofold. On the one hand, it seems to be very interesting to check whether there is empirical evidence for correlation between decision makers. On the other hand, from a more technical point of view, ignoring existing dependencies of other people's decisions could lead to biased parameter estimates (cf. Anselin and Griffith (1988)) or to biased estimates of standard errors (McMillen (1995), for the probit case).

But, why exactly should a decision maker pay attention to the decisions that other individuals have reached? Such social interactions are rational if the utility expected to result from a particular decision is raised as obeying other individuals means an enlargement of the own relevant information set (Brock and Durlauf (1995)). Besley and Case (1995) provide an example of how asymmetric information can lead to such mechanisms: voters lack information which would enable them to judge the state or county government on an absolute scale. Therefore, they tend to make comparisons between incumbents and end up with forcing them into a yardstick competition, using a relative measure instead of an absolute one.

Topa (1997) hints to the point that valuable information may be communicated secretly in social networks, which may result in correlating decisions. Labor market theory provides yet another reason: In some situations, not one's absolute utility may be considered as relevant, but his relative position. In this line is the arguing of the fair wage-effort hypothesis (see e.g. Akerlof and Yellen (1990)): workers want to be paid at a 'fair' level. This puts pressure on managers to conform their wage decision to the level of comparable, i.e. social neighboring, firms. Labor market related behavior of investment managers is as well blamed for the so called herd behavior of professional investors on financial markets, according to Scharfstein and Stein (1990). As 'it is better for reputation to fail conventionally than to succeed unconventionally', few incentives are given not to follow the herd. However, there's a general human tendency to conform to others' behavior which can be perfectly rational, as the society rewards not deviating from accepted norms (Bernheim (1994)).

The wish for conformity may be even higher when there are only two alternatives, leaving no room to opt for a compromise. However, in many situations individuals are faced with discrete choices, e.g. in which foreign country to build a new plant, whether to do an IPO or to start an innovation project are some examples of binary choices.

In this paper, we would like to show how such influences can be considered when modelling and estimating decision behavior. As we assume that such influence are especially prevalent for discrete choice decisions, we concentrate on models of binary and multiple choice. Therefore, we suggest a model which would identify social neighbors - that's what we call decision makers who have to solve similar decision problems - and would estimate the influence neighbors have on each other. The neighborhood defining characteristics are used as dimensions of an abstract space in which the individuals are located. In contrast to the Euclidean assumption commonly used, the space's design which we suggest is more general in the sense that the dimensions are not restricted to be orthogonal. This allows to correct for possible correlations between the social characteristics. Estimating such models poses a sample size problem: in small samples, estimation quality is

poor, whereas the size of the distance matrix can be challenging, as it increases by the square of the sample size. The sample size problem is in the center of our Monte Carlo study.

We apply the suggested methodology to German innovation decision data. Decisions whether about innovation efforts are highly likely to be influenced by the behavior of firms that are competitors on the output and/or on a factor market. The case of product innovations perfectly fits to our considerations, as deciders can rather easy observe product innovations of their competitors when they have been realized. In contrast, they only have an imperfect inside about reasons and success of the innovations, and about planned innovations yet to be realized. As deciders cannot learn by observing these more or less secret features, they will be restricted to utilize as an additional information only the pure information of the observed decision to realize a product information. An incentive for mimicing a competitor's decision can also be labor market pressures, described above for financial investment managers, to which firms' managers may be subjects.

The remainder of the paper is organized as follows: The next section will discuss reasons for including neighborhood influences and present the model and the estimation strategy. In section 3, Monte Carlo studies will show how our model works and illustrate the influence the sample size has on the estimates' quality. The application to the innovation decision follows in section 4. Section 5 concludes and provides some additional ideas for research on this topic.

## 2. AN ECONOMETRIC APPROACH TO INCLUDE NEIGHBORHOOD INFLUENCES

**2.1. General considerations.** Technically, the interdependence of decisions leads to correlations between some or all decision makers. Such cross section correlation between observations is a well-known phenomenon in regional science where contiguous or, more general, neighboring spatial units often influence each other; e.g. the unemployment rates of counties are spatially correlated because an individual is not restricted to working in the county he lives in. Therefore,

cross correlation is often called spatial correlation. But this occurrence is not restricted to geographical spaces: observations can be thought of as being located in an abstract space, with certain social characteristics being the dimensions. Then, if observations are neighbors in this abstract space, these observations are said to be similar.

Meanwhile, a broad range of methods exists - at least for continuous dependent variables - which analyze spatial correlation<sup>1</sup>. However, relatively little research has been done on limited dependent variables in the context of spatial correlation<sup>2</sup>. In an innovation adoption framework, Case (1992) presents an estimation strategy for probit models with spatial dependence in the explained decision result. Unfortunately, the estimation scheme is only applicable to the unsatisfactory case of block wise dependence: the influence parameter between two potential adopters is set to one if they are located in the same region, and to zero otherwise. Another feature which is questionable in the context of (innovation) decisions is the assumption that the different individuals make their decision simultaneously. Often, an individual is only able to observe others' decisions after they have been made, so taking the time pattern into consideration is indispensable for many situations. A spatial innovation diffusion process for logit models is presented by Dubin (1995): In her two period model, firms which have to make a decision do not influence themselves simultaneously. Instead, if a firm did not decide to innovate in the first period, it reconsiders whether to innovate in the second period, and then pays attention to the decisions that other firms have already made in the first period.

The binary decision  $y_i$  of firm  $i$  in the second period depends on the latent scalar variable  $y_i^*$

$$(1) \quad y_i = \begin{cases} 1 & , \quad \text{if } y_i^* > 0, \\ 0 & , \quad \text{else,} \end{cases}$$

and the latent variable specified as

$$(2) \quad y_i^* = x_i' \beta + \sum_{j=1}^N \rho_{ij} y_j^{-1} + u_i, \quad i = 1, \dots, N,$$

---

<sup>1</sup>For an introduction, see the textbook of Anselin (1988), or Part II of Florax (1992).

<sup>2</sup>See Poirier and Ruud (1988), McMillen (1992), Pinkse and Slade (1998).

where  $x_i$  is a vector of the explanatory variables associated with firm  $i$ , with  $\beta$  as the corresponding coefficient vector. The outside influence emanates from decisions  $y_j^{-1}$  which all  $j$  firms made in the first period, weighted by the influence  $\rho_{ij}$  which firm  $j$  has on firm  $i$ , and  $\rho_{ij}$  being 0. The logistic error term (assumed to be i.i.d.) is denoted by  $u_i$ .

Hence, Dubin models innovation decisions that depend on the firms' (geographical) distance to prior innovators - more specifically, adopting a new technology becomes more likely the more first period adopters are in the neighborhood. In this framework Dubin makes a restrictive assumption: Only firms who have not yet adopted an innovation in the prior period are allowed to face an innovation decision in the current period, i.e. firms can innovate only in one of the two periods. This refers to an epidemic diffusion modelling which is restricted to one period which is affected by the dispersion.

In the following section, we present a model which is more general in three respects:

First, decisions depend not only on prior decisions of others but also on the *own* prior decision. Herewith, we do not only measure the influence a decision maker may have on himself, but we also avoid that, when estimating, this influence is caught by the neighborhood effect. We assume all decisions made in this first period as exogenously given.

Second, our approach is not restricted to binary choice problems. However, the multiple choice setting even provides the possibility of estimating specific neighborhood impacts, i.e. neighborhood relationships depending also on each single alternative, offering the possibility to deal with sophisticated decision problems.

Third, we provide a suggestion for an advanced solution to the central problem one faces when modelling neighborhood influences: providing an appropriate concept to measure the degree of similarity, i.e. the 'social distance' between individuals. The most convenient strategy is to define a space within which we are able to measure distances between decision makers. Whereas Case (1992) and Dubin (1995) analyze two-dimensional geographic spaces, we allow for multi-dimensional abstract spaces. We suppose that nowadays, geographic distance

becomes less and less important while individuals and firms pay more and more attention to those being in a comparable situation. For similar reasons, Case and Katz (1991) successfully use a concept of social instead of geographic neighborhood. (See Akerlof (1997), for a model of social space.) Griliches (1992) established the innovation space, using certain characteristics of firms as dimensions for the abstract space and measures these abstract distances by using an Euclidean measure. The drawback of using an Euclidean space is the implied assumption of uncorrelatedness of the social characteristics, which corresponds to orthogonal dimensions spanning the space. We develop this concept by dropping this unrealistic assumption of an Euclidean space. Instead, we allow the dimensions of our space to be not orthogonal according to correlations between the characteristics.

**2.2. The basic model of binary choice decisions.** We inspect the decision making process of  $N$  individuals who are confronted with a certain binary choice problem. Their decisions depend on their own characteristics, on the own decision with respect to the same problem in a prior period (which is in contrast to Dubin's approach), and on decisions of the other decision makers in this prior period (for which we use an alternative modelling). In the binary setup, the decision  $y_i$  of decision maker  $i$  depends, as in equation (1), on the latent variable  $y_i^*$  which we model as

$$(3) \quad y_i^* = \alpha + x_i' \beta + y_i^{-1} \gamma + A_i^{-1} + u_i, \quad i = 1, \dots, N,$$

where  $\alpha$  is a scalar constant.  $x_i$  is a  $K \times 1$  vector of  $K$  explanatory variables associated with  $i$ , and  $\beta$  is the corresponding coefficient vector.  $y_i^{-1}$  describes the decision outcome of individual  $i$  in the prior period (i.e., 0 or 1), with the scalar parameter  $\gamma$  measuring the impact this own decision has on the current period.  $A_i^{-1}$  is a scalar reflecting the neighborhood impact of all other individuals on decision maker  $i$ , and  $u_i$  is an i.i.d.  $N(0, \sigma^2)$  error term, where  $E[A_i^{-1} u_i] = 0$  is assumed. We model  $A_i^{-1}$  as the weighted sum of the decisions  $y_j^{-1}$  which the individuals  $j$ , that are  $i$ 's neighbors, made in the prior period. The weights used for  $A_i^{-1}$  depend on the distance  $D_{ij}^*$  between  $i$  and  $j$ , and on parameters

determining the scale  $(a_1, a_0)$  and the decay  $(b_1, b_0)$  of the neighborhood influence

$$(4) \quad A_i^{-1} = \sum_{\substack{j=1 \\ j \neq i}}^N \left\{ [a_1 \cdot \exp(-D_{ij}^* b_1) \cdot y_j^{-1}] + [a_0 \cdot \exp(-D_{ij}^* b_0) \cdot (1 - y_j^{-1})] \right\}.$$

This specification allows for different spatial dependence patterns, conditional on the decision of the particular neighbor. The spatial parameters  $(a_1, a_0, b_1, b_0)$  are, together with  $\alpha, \beta$  and  $\gamma$ , the parameters to be estimated.

In the following we discuss the calculation of the neighborhood impact  $A_i^{-1}$ . A prerequisite for this is to determine the distance  $D_{ij}^*$  between all pairs of decision makers  $i$  and  $j$ . As the distance shall be small if the decisions makers are in similar situations (i.e. the decision makers shall, in this case, be treated as ‘neighbors’), the distance measure has to depend on characteristics that describe the decision makers’ situations. Therefore, we define the  $D \times N$  matrix  $Z$  of (social distance defining) characteristics by

$$Z = \begin{pmatrix} z_{11} & \dots & z_{1N} \\ \vdots & & \vdots \\ z_{D1} & \dots & z_{DN} \end{pmatrix},$$

where  $z_{di}$  denotes the social characteristic  $d$  ( $d = 1, \dots, D$ ) of individual  $i$ . Furthermore, we define a  $D \times 1$  vector  $\zeta_{ij}$  of all the (social) differences between an individual  $i$  and an arbitrary other decision makers  $j$  ( $j \neq i$ ) within the  $D$  characteristics by<sup>3</sup>

$$\zeta_{ij} \equiv \begin{pmatrix} z_{1i} - z_{1j} \\ z_{2i} - z_{2j} \\ \vdots \\ z_{Di} - z_{Dj} \end{pmatrix}.$$

Vector  $\zeta_{ij}$  can be interpreted as being located in a  $D$ -dimensional abstract space, which is spanned by the  $D$  characteristics. The most popular strategy is to measure distances in an Euclidean space where the Euclidean Distance is given by

$$(5) \quad D_{ij} = \sqrt{\zeta_{ij}' \zeta_{ij}}.$$

---

<sup>3</sup>For ease of notation we assume that the particular social characteristics have a mean equal zero.



The assumption of an Euclidean space is, in this context, equivalent to assuming uncorrelatedness between the characteristics as well as equal variances, i.e. the correlation matrix of the social characteristics is assumed to be the identity matrix. This is a serious restriction to the general case of arbitrary correlation between the characteristics: In most applications, like for innovation decisions, such characteristics will be correlated, e.g. the age of a firm and its size. Furthermore, there may be information concerning the size in terms of the number of employees and the turnover as well. If the correlation will be considered, the information of both characteristics can be used without overemphasizing the size aspect. Therefore, we propose a transformation of the differences  $\zeta_{ij}$  given by

$$(6) \quad \zeta_{ij}^* \equiv \sqrt{2N}P\zeta_{ij},$$

where  $P$  is obtained by the Cholesky factorization of  $ZZ'$ , i.e. <sup>4</sup>

$$(7) \quad P'P = (ZZ')^{-1}.$$

Nevertheless, the characteristics of different decision makers are still assumed to be independent. The transformation of  $\zeta_{ij}$  into  $\zeta_{ij}^*$  is, of course, equivalent to transforming  $Z$  by  $Z^* \equiv \sqrt{N}ZP$ , and calculating  $\zeta_{ij}^*$  based on  $Z^*$ .

This modification procedure ensures that the modified differences in  $\zeta_{ij}^*$  are orthogonal, thus, they now suit to be transformed into distances by the Euclidean measure. Hence, our modified distance measure is defined by

$$(8) \quad D_{ij}^* = \sqrt{\zeta_{ij}^{*'}\zeta_{ij}^*}.$$

This concept provides, in view of the stochastic properties of the characteristics, a more general and more realistic measure of distances in an abstract space. If the characteristics are uncorrelated and the variances of all characteristics are equal,  $P\sqrt{N}$  is the identity matrix, thus  $\zeta_{ij}^* = \zeta_{ij}$ . Hence, our modified distance measure nests the Euclidean distance measure as a special case.

---

<sup>4</sup>If the matrix  $ZZ'$  does not have full rank the Cholesky factors  $P$  have to be computed on the generalized Moore-Penrose inverse. Hence, if one characteristic vector of the space can be written as a linear combination of other vectors the number of the dimensions of the space will be reduced by one.

After determining the distance between all pairs of decision makers, the next step consists in defining an adequate distance decay function (DDF) which translates the social distance into a value expressing the strength of the neighborhood impact. As this influence of  $j$  on  $i$  shall diminish with distance, a DDF  $\rho = \rho(D_{ij}^*)$  has to be defined which (i) is decreasing in the argument, (ii) returns positive values for positive arguments and (iii) returns a finite positive value for zero. A functional form fulfilling this requirement is the DDF

$$(9) \quad \rho(D_{ij}^*) = a \exp(-D_{ij}^* \cdot b),$$

which offers the advantage that two parameters are enough to describe a rather wide range of different decay patterns: The parameter  $a$  measures the strength of the neighborhood influence on the decision outcome, whereas the parameter  $b$  allows us to investigate the rate by which the impact of prior decisions attenuates with the distance. Obviously, the parameter  $b$  is only identified if  $a \neq 0$ . If the parameters  $a$  and  $b$  are jointly significant, empirical evidence for neighborhood influences is found.<sup>5</sup>

A reasonable extension of the model is motivated by the possibility that makers of different decisions may have a different distance decay. Such differences are particularly likely if one of the both alternatives is opted for rarely, so decisions for this alternative are noticed carefully with little respect to the distance to the respective decision maker. For example, one observed such asymmetric patterns in the early days of the internet, when the launch of any firm's web site was exceedingly noticed. Therefore, introducing two different DDFs  $\rho_1$  and  $\rho_0$  (which are associated to different parameters  $a_1$  and  $b_1$  and, resp.,  $a_0$  and  $b_0$ ) may be sensible for certain circumstances.

The neighborhood impact of  $j$  on the decision outcome of  $i$  is then  $j$ 's weighted decision outcome

$$(10) \quad \eta_{ij} = [\rho_1(D_{ij}^*) \cdot y_j^{-1}] + [\rho_0(D_{ij}^*) \cdot (1 - y_j^{-1})].$$

---

<sup>5</sup>Of course, there are many other functional forms suitable as DDF. For the Monte Carlo as well as for the application study, we also employed alternative functions which failed to give evidence of any superiority.

Aggregating the spatial impacts  $\eta_{ij}$  of all units  $j \neq i$  leads to

$$A_i^{-1} = \sum_{\substack{j=1 \\ j \neq i}}^N \eta_{ij},$$

which is equivalent to equation (4).

Neglecting neighborhood dependencies, i.e., dropping  $A_i^{-1}$  in equation (3), corresponds to the problem of omitted variables. Thus, the coefficient vector  $\beta$  would be estimated with a bias which depends on the correlations between the explanatory variables  $x$  on the one hand and the distances within the characteristics  $Z$  and prior period decisions  $y_j^{-1}$  on the other hand. Additionally, the estimated standard errors of  $\beta$  would be wrong.

**2.3. Multiple Choice.** Often decision makers have to decide not only between ‘yes’ or ‘no’ but between several different alternatives. Therefore, a more general specification is obtained by regarding more than two alternatives. For this purpose, the binary decision model of subsection 2.2 can be generalized by including spatial dependencies in multinomial models. By using alternative specific spatial parameters, it is possible to estimate the particular impact that decision makers, who chose in the prior period a certain alternative  $r$ , have on individuals choosing between  $S$  alternatives in the next period.

The general setting of the multiple choice model is according to the basic binary choice model: An individual  $i$  can decide among  $S$  alternatives, depending again on his own characteristics, and on the (exogenously given) decisions made by him and his social neighbors  $j$  in a certain prior period. Each alternative  $s$  provides an outcome (utility)  $y_{is}^*$ ,  $s = 1, \dots, S$ , which is not observable. The decision maker chooses the alternative  $s$  if  $y_{is}^* > y_{is'}^*$ ,  $\forall s' \neq s$ . Then, the observable variable  $y_i$  takes the value  $y_i = s$ . By separating the impact of the own prior decision we define the latent model associated with the choice of  $s$  by

$$(11) \quad y_{is}^* = x_i' \beta_s + \sum_{r=1}^R y_{ir}^{-1} \gamma_{rs} + A_{is}^{-1} + u_{is}, \quad i = 1, \dots, N;$$

where  $y_{ir}^{-1} = \mathbb{1}_{(y_i^{-1} = r)}$ .  $x_i$  denotes, as in section 2.2, a vector of explanatory variables associated with decision maker  $i$  who chooses alternative  $s$  in the

current period, with  $\beta_s$  as the corresponding coefficient vector.  $\gamma_{rs}$  measures the strength of the own prior decision's impact if the decision maker chose alternative  $r$  ( $r = 1, \dots, R$ ) in the prior period. Thus, this specification allows quantifying the particular impacts of the own prior decisions depending on the chosen alternatives in the two periods. In the prior period, the number  $R$  of alternatives may differ from the alternatives' number  $S$  in the current period. For the error terms  $u_i \equiv (u_{i1}, \dots, u_{iS})'$  we assume  $u_i \sim \text{i.i.d. } N(0, \Sigma)$  and  $E[A_{is}^{-1}u_{is}] = 0$ .  $A_{is}^{-1}$  denotes the aggregated spatial impacts of all prior decisions on a decision maker  $i$  choosing the alternative  $s$ :

$$(12) \quad A_{is}^{-1} = \sum_{\substack{j=1 \\ j \neq i}}^N \left[ \sum_{r=1}^R a_{rs} \exp(-D_{ij}^* b_{rs}) y_{jr}^{-1} \right],$$

being a straightforward generalization of (4), with  $a_{rs}$  and  $b_{rs}$  measuring scale and pattern of the influence that prior decision makers who chose alternative  $r$  have on decision makers choosing alternative  $s$ . As before, these spatial parameters are to be estimated.

Again, an estimation of  $\beta_s$  that ignores the impact of prior decisions will generally be biased.

### 3. MONTE CARLO STUDIES

The inclusion of a spatial distance matrix into discrete choice models leads to highly nonlinear models which poses the question after the small sample properties of our estimation scheme. Especially with regard to the application in section 4, we want to gain insight into how the quality of estimates depends on different sample sizes.

We analyze the small sample properties for a correctly specified binary choice model defined according to equations (3) and (4) which is estimated by Maximum Likelihood. For the exogenously given decision outcomes  $y_i^{-1}$  from the prior period we use random draws  $v_i$  from a uniformly  $U[0, 1]$  distribution coded by  $y_i^{-1} = \mathbb{1}(v_i > 0.5)$ . We choose a parameter constellation which ensures that the ratio between decisions of both outcomes (0 and 1) is about balanced. We fixed the parameters as follows: constant  $\alpha = 1$ , one slope parameter  $\beta = -1$ , and

the parameter for the impact of the own prior period decisions is  $\gamma = 1$ . The spatial parameters are chosen to be  $a_1 = 1$  and  $a_0 = -3$  as scaling and  $b_1 = 3$  and  $b_0 = 5$  as decay parameters. The social space is spanned by two orthogonal vectors obtained by random draws of the  $U[0, 1]$  distribution. As we want to concentrate on the estimation quality dependent subject to the sample size, we do not utilize the correction described in 6. The error term is drawn from a standard normal distribution, while the exogenous variables  $x_i$  are sampled from the uniformly  $U[0, 1]$  distribution.

We analyze eight sample sizes from 50 to 400 with 1,000 replications each. The results are presented by means of box plots in appendix 6.1. There, the figures 1 till 5 show the empirical distributions of the respective parameters depending on the sample size. The horizontal line within each box indicates the empirical distribution's median. The box itself is defined by the upper and lower quartile. Adding three halves of the interquartile range into both directions from the box yields the position of the so-called fences. Adjacent values, which lie out of the box-fences-formation, are denoted by a small circle.<sup>6</sup> The respective true value is located in the middle of each figure and marked by a horizontal line. Note that the scales of the y-axes differ. Concerning the spatial parameters, we only present the two ones which describe the influence of the former decisions  $y_i^{-1} = 1$ , as the estimates of the two other parameters with respect to  $y_i^{-1} = 0$  were qualitatively identical.

A general finding of our study is that the challenge, which seven parameters and a rather complicated likelihood mean at least for a sample size of 50, has been met: there are no convergence problems to notice and no sensitivity to the start values even for this smallest sample.

For all parameters, the variance of the estimates' empirical distributions reduces remarkably over the sample size range. For  $n = 50$ , there are, for any parameter, a handful of extreme outliers which can not be drawn into the figure. But for rising sample size, the variance converges at high speed, and for  $n = 400$ ,

---

<sup>6</sup>See Tukey (1977) for a detailed presentation of the box plot concept.

the reduction of the estimator's variance achieved by additional observations is only considerable for the two spatial parameters.

Again, all five types of parameters show a clear small sample bias. This is no shift of the empirical distribution but an asymmetry which leads to overestimation in terms of the absolute values. However, this bias disappears to different degrees for rising sample sizes and does not vanish completely for any parameter under consideration. It reveals to be strongest for the own impact parameter  $\gamma$  and the extremely asymmetric spatial scaling parameter  $a_1$ , while the decay parameter's  $b_1$  estimates are the least biased ones.

As a result, despite some shortcomings, the Maximum Likelihood estimation of our probit model proves to be possible and meaningful. The shortcomings are, for all kinds of parameters, a rather high variance when the sample size does not exceed 100, and a considerable bias which does not vanish even for some hundreds observations.

#### 4. APPLICATION TO INNOVATION DATA

Research on R&D and innovation activities is familiar with circumstances in which decisions of a firm - namely to aim for and to adopt a new technology - do not only depend on the firm's own situation (see e.g. Geroski (1995) or Encaoua, Hall, Laisney, and Mairesse (1998)). Reasons for that are mainly, but not only, effects of networks and of standardization (for a review, see e.g. Katz and Shapiro (1985)) and, of course, of research spillovers (the broad discussion about this subject has been established by Griliches (1992)). Further reasons, quite in line with those we discuss in our introduction, are compiled by Baptista (2000) and Manski (2000). Consequently, we apply our model to data of product innovation activities of German firms because the features of this decision - to innovate or not - suits well to our general model setting: It is a binary decision, the degree of uncertainty is rather high which makes the information how others decided more valuable, the decision makers are under the labor market's conformity pressure described in the introduction, and the competition between firms urges to keep pace with the technological progress of one's rivals.

The data we use is the first wave of the Mannheim Innovation Panel (MIP), provided by the Centre for European Economic Research (ZEW, Mannheim; see Harhoff and Licht (1994) for details on the data composition.). The MIP is a survey of approximately 3000 German firms which has been annually collected since 1993. Its questionnaire follows the guidelines of the OSLO-manual (OECD (1992)) and contains detailed information on potential determinants of innovative activity. In contrast to our (implicit) assumption, our data does not provide information about all firms potentially influencing the firm under consideration. Having this drawback in mind, we assume that the sample units, which have been randomly selected from the total population when the MIP has been designed, can be used to predict all outside influence a firm receives. Because the MIP provides binary information in each of the preceding three years as to whether an innovation has taken place, we apply the binary choice framework of section 2.2. By using the outcomes from 1990 and 1992, we analyze a sample of all firms in the MIP for which the relevant information is available, in total 1380 observations. Table 1 (section 6.2) shows the number of the adopted product innovations in the years 1990 and 1992. We use the entirety of the MIP data instead of single industries' data because doing the latter would lead to sample sizes of about 100 or less. As the Monte Carlo study in the previous chapter shows, such samples are too small for reliable estimations.

It would be beyond the scope of our paper to derive an estimation equation from a structural approach. For our reduced form approach, we use two groups of variables as determinants of the innovation decisions. While the first one reflects the ability of a firm to cope with the effort necessary for a successful innovation, the second group provides information on the subjective view managers have on their firms' situation. The first group of 'hard' variables encompasses the employees' number, the square of this number, and the ratio of own capital to total sales. The latter serves as an indicator for the firm's financial strength. The 'soft' variables are two expectations a firm's manager will have concerning the relevant product market: the developments of the demand for their products on the one hand, and the competition on this market on the other hand.

In order to measure the distance (i.e., to define neighborhood) between the firms one has to determine suitable dimensions for the abstract space. We utilize six variables as these characteristics composing the matrix  $Z$ , which, again, can be divided in ‘hard’ variables that describe the firm’s situation, and ‘soft’ variables that allow to see into the managers’ motivation. For the first group of variables, we utilize the firm’s age, the share its exports have in its sales, the sales share of its most important product, and whether it is located in Eastern Germany. For the latter group, we use two questions in which firms have been asked to weigh the importance of possible objectives for their innovation decision on a scale from 1 (no importance) to 5 (high importance). We employ ‘Decreasing the impact on the environment’ and ‘Lowering the energy consumption’. Table 2 shows the correlation matrices of the chosen six social characteristics. Considerable correlations between particular characteristics point out the necessity to correct for it by using the modified distance measure as defined in equation (6).

The results of the Maximum Likelihood estimations are presented in table 3. In particular, we employ four specifications.

The first specification is a common probit without any own or neighborhood influence from the previous period. The estimation yields coefficients with the expected sign which are significant at least on the 10%-level, with the exception of own capital per sale. The size of a firm, expressed by the number of employees, has a decreasingly positive effect on the innovation probability, with a probability maximizing firm size of roughly 100 employees. The higher the expected demand in the corresponding product markets the higher the innovation probability, which is little surprising. The tendency to innovate also rises with the competition in the product market a firm expects; supposable, a product innovation is, at least for some firms, an important strategic move in the market competition. Remarkably, the estimated coefficients for the two ‘soft’ variables expected demand and the expected competition prove to be very stable within the other three specifications (b) till (d) which is not similarly true for the three ‘hard’ variables.

More precisely, the inclusion of the own prior decision in specification (b) affects the firm size variables estimates and makes them insignificant. This might be



due to the mostly low changes of employment figures within two years, while the opinion concerning the demand and the competitors' behavior may alter more quickly. The variable measuring the own prior effect is strongly significant positive, as expected.

The third specification (c) includes the neighborhood dependence effect, allowing for positive influence exerted by innovators, and for negative influence exerted by the non-innovators, assuming an identical decay for both kinds of influence. The decay parameter  $b$  is strongly significant, which expresses the higher relevance neighbors have on a decision. The two scale parameters, being both significant on the 5 %-level, clearly reveal the assumed influence pattern. The Wald test on joint significance of the three spatial parameters is significant, as well.

The full model, which allows for different decays for innovators' and non-innovators' influences, is presented in the last specification (d). It states the results found before. The scale parameters only change mildly and keep their significance level. The two values for the decays are roughly of the same size. As all four spatial parameters are significant at least on the 5 %-level, the Wald statistic for their joint significance is, again, persuasively significant.

## 5. CONCLUSIONS AND OUTLOOK

In this paper, we model and estimate discrete choice problems of individuals under the assumption that decisions which others have made in similar situations influence the individual decision maker. According to the distinction described e.g. by Manski (1993), these outside influences are interactions of an endogenous kind, as the propensity for a certain decision outcome depends on the *outcomes* of other decision makers who are relevant due to their social similarity. To measure the degree of social similarity between individuals we define an abstract space so that agents with similar decision problems are neighbors in this space. Our approach takes correlations between individuals' characteristics, which define the dimensions of the abstract space, into account. We present Monte Carlo studies for a binary choice model which show that even a sample size of  $n = 50$  can yield

meaningful results. Nevertheless, unbiasedness and variance of the empirical distribution of the estimates heavily improve for a sample of, say,  $n = 200$ . An application to German innovation data, which estimates the probability to decide for a product innovation, reveals strong evidence for the existence of the described neighborhood influence.

The next research steps should explore certain aspects of the presented approach. A promising possibility seems to be the widening of the time horizon: with panel data, the innovation decision pattern over several periods could be described. Another aspect worth examining is the design of the distance decay function. For concrete applications, it remains demanding to detect not only the degree of importance which an outside influence has on a decision maker but especially the factors on which this influence depends. The crucial point is, technically, to fix how the distances between decision makers can be properly determined, which will, in most cases, raise the question as to which factors express similarity between the agents' decisions. Against that background, learning about social interactions could benefit from data containing information on subjective views of the decision makers, as e.g. Manski (2000) emphasizes.

## REFERENCES

- AKERLOF, G. (1997): "Social Distance and Social Decisions," *Econometrica*, 65, 1005–1027.
- AKERLOF, G., AND J. YELLEN (1990): "The Fair Wage-Effort Hypothesis and Unemployment," *Quarterly Journal of Economics*, 105, 255–283.
- ANSELIN, L. (1988): "Spatial Econometrics: Methods and Models," *Kluwer, Dordrecht*.
- ANSELIN, L., AND D. GRIFFITH (1988): "Do Spatial Effects Really Matter in Regression Analysis?," *Papers of the Regional Science Association*, 65, 11–34.
- BAPTISTA, R. (2000): "Do Innovations Diffuse Faster Within Geographical Clusters?," *International Journal of Industrial Organization*, 18, 515–535.
- BERNHEIM, B. (1994): "A Theory of Conformity," *Journal of Political Economy*, 102, 841–877.
- BESLEY, T., AND A. CASE (1995): "Incumbent Behaviour: Vote-Seeking, Tax-Setting, and Yardstick Competition," *American Economic Review*, 85, 25–45.
- BROCK, W., AND S. DURLAUF (1995): "Discrete Choice with Social Interaction I: Theory," Discussion Paper 5291, NBER Working Papers.
- CASE, A. (1992): "Neighborhood Influence and Technological Change," *Regional Science and Urban Economics*, 22, 491–508.
- CASE, A., AND L. KATZ (1991): "The Company You Keep: The Effects of Family and Neighborhood on Disadvantaged Youths," Discussion Paper 3705, NBER Working Papers.
- DUBIN, R. (1995): "Estimating Logit Models with Spatial Dependence," *L. Anselin and R. Florax, "New Directions in Spatial Econometrics"*, Springer, Berlin, pp. 229–242.
- ENCAOUA, D., B. HALL, F. LAISNEY, AND J. MAIRESSE (1998): "The Economics and Econometrics of Innovation," *Annales d'Economie et de Statistique*, 49/50, 27–51.

- FLORAX, R. (1992): "The University: A Regional Booster?," *Avebury, Aldershot*, pp. 189–228.
- GEROSKI, P. A. (1995): "Markets for Technology: Knowledge, Innovation and Appropriability," in *Handbook of the Economics of Innovation and Technological Change*, ed. by P. Stoneman, chap. 4, pp. 90–131. Blackwell, Oxford.
- GRILICHES, Z. (1992): "The Search for R&D Spillovers," *Scandinavian Journal of Economics*, 94(supplement), 29–47.
- HARHOFF, D., AND G. LICHT (1994): "Das Mannheimer Innovationspanel," *U. Hochmuth and J. Wagner, "Firmenpanelstudien in Deutschland", Tuebinger Volkswirtschaftliche Schriften, Tuebingen*, pp. 255–284.
- KATZ, M., AND C. SHAPIRO (1985): "Network Externalities, Competition and Compatibility," *American Economic Review*, 75, 424–440.
- MANSKI, C. F. (1993): "Identification of Endogenous Social Effects: The Reflection Problem," *Review of Economic Studies*, 60, 531–542.
- (2000): "Economic Analysis of Social Interactions," Discussion Paper 7580, NBER Working Papers.
- McMILLEN, D. (1992): "Probit with Spatial Autocorrelation," *Journal of Regional Science*, 32, 335–348.
- (1995): "Spatial Effects in Probit Models," *L. Anselin and R. Florax, "New Directions in Spatial Econometrics", Springer, Berlin*, pp. 189–228.
- OECD (1992): "Proposed Guidelines for Collecting and Interpreting Technological Innovation Data - OSLO Manual," *OECD/DG (92)*, 26.
- PINKSE, J., AND M. SLADE (1998): "Contracting in Space: An Application of Spatial Statistics to Discrete-Choice Models," *Journal of Econometrics*, 85, 125–154.
- POIRIER, D., AND P. RUUD (1988): "Probit with Dependent Observations," *Review of Economic Studies*, 55, 593–614.
- SCHARFSTEIN, D., AND J. STEIN (1990): "Herd Behaviour and Investment," *Quarterly Journal of Economics*, 105, 255–283.
- TOPA, G. (1997): "Social Interactions, Local Spillovers and Unemployment," Discussion paper, Department of Economics, New York University.
- TUKEY, J. W. (1977): *Exploratory Data Analysis*. Addison-Wesley, Reading, MA.

## 6. APPENDIX

**6.1. Monte Carlo Simulation Results.** The following box plot figures show the empirical distributions of the estimated parameters for different sample sizes, which resulted from probit estimations with 1,000 replications each. See section 3 for a detailed description of the study's design.

FIGURE 1. Empirical distributions of constant's estimates.

Estimation of the constant  $\alpha$ .  
(True value:  $\alpha = 1$ )

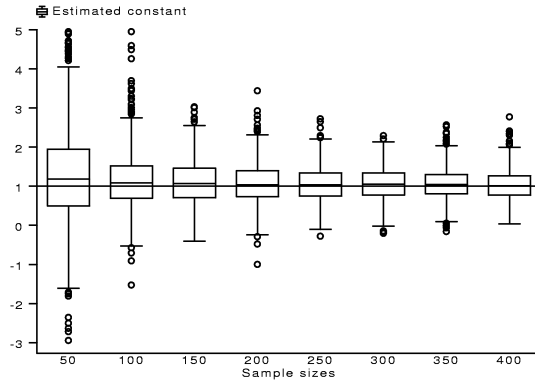


FIGURE 2. Empirical distributions of slope parameter's estimates.

Estimation of the slope parameter  $\beta$ .  
(True value:  $\beta = -1$ )

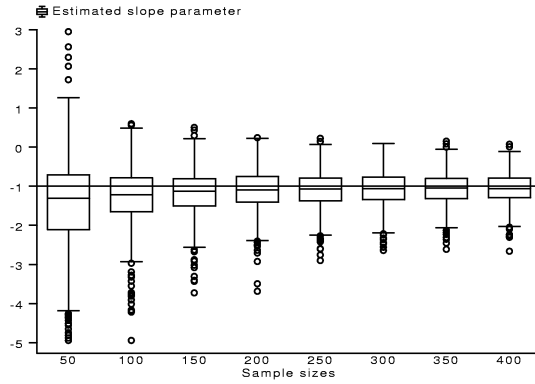


FIGURE 3. Empirical distributions of own impact parameter's estimates.

Estimation of the parameter  $\gamma$  which reflects the impact from the own prior period's decision. (True value:  $\gamma = 1$ )

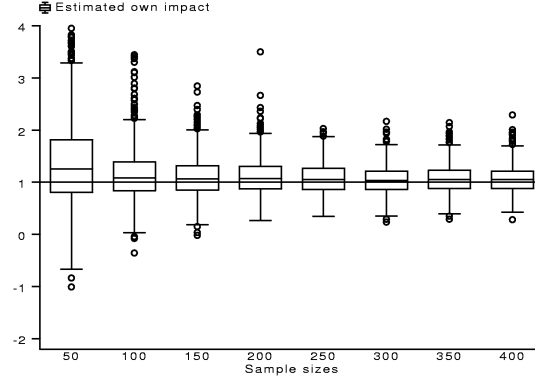


FIGURE 4. Empirical distributions of scaling parameter's estimates.

Estimation of the spatial scaling parameter  $a_1$  which expresses the size of that neighbors' influence who had decided positively (i.e.,  $y_i^{-1} = 1$ ) in the prior period. (True value:  $a_1 = 1$ )

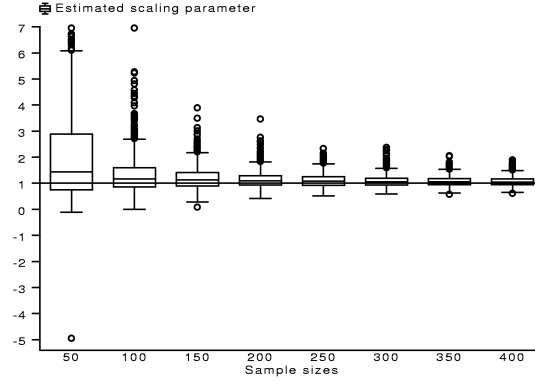
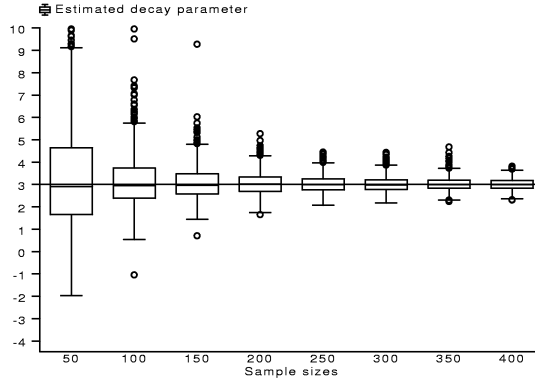


FIGURE 5. Empirical distributions of decay parameter's estimates.

Estimation of the spatial decay parameter  $b_1$  which expresses the decay of that neighbors' influence who had decided positively (i.e.,  $y_i^{-1} = 1$ ) in the prior period. (True value:  $b_1 = 3$ )



## 6.2. Estimation of Product Innovation Decisions in Germany.

TABLE 1. Number of product innovators in the years 1990 and 1992. Based on data of the Mannheimer Innovationspanel.

		1992		
		No	Yes	
1990	No	186	375	561
	Yes	69	750	819
		255	1125	1380

TABLE 2. Correlation matrices of social characteristics. The characteristics are denoted by

A: importance of objection ‘less impact on environment’ and of

B: ‘lower energy consumption’ for the own innovation decision,

C: firm age,

D: share of exports,

E: located in eastern Germany,

F: sales share of most important product.

	A	B	C	D	E	F
A	1.0000					
B	0.3023	1.0000				
C	0.0337	-0.0350	1.0000			
D	0.0265	-0.0382	0.1765	1.0000		
E	-0.0507	0.1572	-0.4193	-0.2953	1.0000	
F	-0.0021	0.0552	-0.0355	0.0030	0.0353	1.0000

TABLE 3. Estimation results of product innovation decisions

$n = 1380$	(a)	(b)	(c)	(d)
<i>Mean Variables</i>				
Constant	0.2758*	-0.3115*	-0.4033**	-0.4090**
Employment	0.0957*	0.0420	0.0498	0.0512
(Employment) <sup>2</sup>	-0.0005*	-0.0002	-0.0002	-0.0002
Own Capital per Sales	-0.0123	0.0315	0.0263	0.0265
Expected Demand	0.0924**	0.1350***	0.1251***	0.1232***
Expected Competition	0.0682*	0.0693*	0.0699*	0.0690*
<i>Own Impact</i>				
$\alpha$	-	0.9597***	0.9677***	0.9732***
<i>Spatial Parameters</i>				
$a_1$	-	-	0.2165**	0.1934**
$a_0$	-	-	-0.1193**	-0.1789**
$b_1$	-	-	3.5275***	3.5050***
$b_0$	-	-	4.1308***	4.1308***
$\chi^2(3)$ resp. $\chi^2(4)$	-	-	59.40***	72.29***
log Likelihood	-651.96	-585.01	-581.72	-581.41

1380 firms of the MIP sample 1990/1992. Dependent variable: Product innovation realized in 1992 (yes/no). Inference based on robust standard errors. Wald test for joint significance of the three resp. four spatial parameters.

\*\*\* : significant on the 1%-level

\*\* : significant on the 5%-level

\* : significant on the 10%-level

no star : no significance